

Bus network structure and mobility pattern: A monocentric analytical approach on a grid street layout

Hugo Badia ^{a,*}, Miquel Estrada ^a, Francesc Robusté ^a

^a Barcelona Tech, Technical University of Catalonia (UPC), C/ Jordi Girona 1-3, Building B1, Office 006, North Campus, Barcelona 08034, Spain

* Corresponding author. Tel.: +34 934137667; fax: +34 934137675.

E-mail addresses: hugo.badia@upc.edu (H. Badia), miquel.estrada@upc.edu (M. Estrada), f.robuste@upc.edu (F. Robusté).

Abstract

This study discusses which transit network structure is the best option to serve urban mobility. As a consequence of the evolution of urban form, cities have undergone a process of dispersion of their activities that has caused a change in mobility needs in the last few decades. Mobility networks and services should progressively adapt to the new demand patterns, especially the bus transit network, which has more flexibility to absorb the changes. We compare four base transit network structures: a radial scheme, a direct trip-based network, and a transfer-based system by means of either a complete grid or a hybrid structure. An analytical model is used to estimate the behavior of these structures for idealized monocentric mobility patterns with several degrees of concentration. The purpose is to determine the right range of situations for the applicability of each bus transit structure, and to determine guidelines about the transit network planning process. It turns out that the best structure is not always the same, and depends on the mobility spatial pattern. A radial network is the best alternative in very concentrated cities; however, a direct trip-based system is more suitable for intermediate degrees of dispersion. A transfer-based structure is the best option when the activities are more decentralized. Nevertheless, the decentralization degree that justifies a specific transit structure is not constant. This degree depends on the characteristics of the city, transport technology and users.

Keywords: Transit network design; bus network structure; hybrid network; public transportation; urban mobility.

1. Introduction

Transportation has a relevant impact on the economy and also is a key issue with regard to social, political and environmental aspects. The economic competitiveness of a city and its social equity might be determined by its transport supply. The convenience of transit systems versus cars in urban areas is clearly accepted. However, a good design of these transit systems is essential to guarantee their efficiency and effectiveness in order to satisfy mobility necessities. Among the different aspects that one can consider for a well-designed transit network, its network structure stands out, i.e., how the different transit lines that compose the system are arranged over the city. This structure should be the most suitable with regard to transit system cost and level of service to satisfy urban mobility patterns.

These urban mobility patterns have been evolving at the same time as urban spatial structures. Rodrigue et al. (2006) summarized this evolution in three phases: the first is a highly centralized scenario, where the major part of activities takes place in a central business district (CBD) or city center; the second is an intermediate phase, when some of these activities start to scatter over adjacent areas of the CBD; and the final phase is a dispersed urban form, where a high number of activities is relocated in new peripheral areas far away from the CBD. Anas et al. (1998) presented a similar evolution of the city in the last two centuries from a centralized to a decentralized scenario due to changes in technology, telecommunications and transportation. From the so-called *nineteenth century city*, characterized by a compact core surrounded by residential areas, CBDs have expanded. At first, activities from the center spread outward around it; afterwards, new poles of activities appear, forming subcenters or, on a metropolitan scale, edge cities. Initially, these new poles have a clustered form, and with time, they also reach a scenario of dispersion.

American cities' forms have been studied in a more exhaustive way than European or Asian cities. Lee (2007) identified three distinctive patterns in American urban areas: (i) monocentric cities, where the urban core remains stable and its decentralization takes place in its adjacent areas; (ii) polycentric cities,

where those activities that are decentralized are concentrated in suburban centers at the same time; and (iii) dispersed cities, characterized by the absence of a clear structure, where decentralized activities from the CBD are not regrouped in other centers. The most generalized pattern for European cities is the first of these patterns, since traditional city centers remain the most significant pole of activities, and new poles are usually dependent on them.

On the other hand, Bertaud (2004) summarized the global trend in the evolution of urban spatial structures in two types of cities: a monocentric city where the traditional CBD was still predominant, although it progressively expanded, and a polycentric city where no predominant centers existed. In addition, the author suggested that the monocentric structure was predominant even in polycentric cities. Perfect polycentrism, as an agglomeration of autonomous urban villages that conform a large city, does not exist in the real world. New poles of demand attracted trips from everywhere in the city, but these were less relevant than the traditional CBD, forming the mono-polycentric model defined in Bertaud (2004). Craig and Ng (2001) showed the decreasing slope of the employment densities from the CBD to the periphery. The decreasing tendency was broken when new centralities appeared. These presented higher densities compared to their surroundings but not globally. The same result was obtained by a spatial econometric analysis in Pereira et al. (2013), where most cities presented a cluster of high demand around the CBD and clusters of low demand in the most external periphery. If some outliers appeared far away from the city center, these only represented locally high values.

Due to this process of decentralization, the vector of displacements changes from a centripetal pattern, where all trips have their origin or destination in the city center, to decentralized scenarios, where a high percentage of trips connect peripheral areas without depending on the CBD. Given this evolution, transit systems should respond to the adaptation of their supply. In a centralized urban form, the best solution is evidently a radial structure; however, when this scenario has been overcome, the best way to serve the demand has been discussed. There are two main strategies: (i) a direct trip-based network and (ii) other transfer-based network, called by some authors a multi-destination system. The former maintains its old radial lines and is complemented by new lines that connect new poles of demand to other areas of the city. In this system, each line works independently of the others, without a clear way as the network operates. Its final objective is to satisfy the highest percentage of displacements by means of direct trips. The latter removes the initial radial lines in order to develop a new network characterized by a simple structure, where transferring is an essential step of the transit chain (Nielsen et al., 2005).

The question is when and where each type of system is the most suitable to serve demand. Thompson (1977) made a simple comparison between four different structures: radial, ubiquitous (the highest degree of a direct trip-based system, where there is a direct line from each origin to any destination), grid and timed transfer (these latter two represent two different transfer-based schemes). The main conclusion was that a transfer-based system offers a better service for peripheral trips than the traditional radial scheme, and, at the same time, it is cheaper than a ubiquitous scheme. As a consequence, in a process of city decentralization, these alternative structures are more suitable. However, the lack of analysis of a direct trip-based system concentrated around the main attraction poles leaves the door open for a more profound comparison. Newell (1979) argued that a central area, where the major part of activities is concentrated, is generally found in cities due to economies of scale. Therefore, the transit system has to focus on that space. For this reason, he maintained the idea that a radial system concentrated around a central corridor would be the best strategy.

Jara-Díaz and Gschwender (2003) studied the same problem in a simple spatial system, only composed of five nodes that form a cross. Two different strategies to serve the demand are compared: direct trips by means of lines that connect the different pairs O/D demanded, or by only two perpendicular corridors with a transfer point at their intersection. Their conclusion was that working with two corridors became more convenient with the demand disparity, and it is conditioned by the level of demand and unit user and agency costs. However, they did not consider some important aspects in the analysis, such as the spatial coverage or the trip length, and did not identify the degree of dispersion from which a transfer-based network is better. An extension of the analysis to a two-dimension network would be necessary.

Other contributions have justified the convenience of transfer-based transit systems by means of analyzing real networks. Thompson and Matoff (2003) and Brown and Thompson (2008) justified the change of transit system approach to transfer-based by comparing the supply, ridership, efficiency and effectiveness of different American metropolitan areas. They concluded that those that had implemented a

transfer-based network had a better performance on most measures, and had not lost productivity. Mees (2000) made an exhaustive comparison about urban form and transit service between two cities, Melbourne and Toronto. The main conclusion was that the main reason why ridership in Toronto was higher than Melbourne was not related to the characteristics of their urban form, which were similar, but rather to the transit service orientation and its consequences. Toronto presented a grid transit network structure following its street pattern. Other comparisons between Boston and Toronto (Schimek, 1997) or Broward (Florida) and Tarrant (Texas), made by Brown and Thompson (2012), gave similar conclusions.

While transfer-based networks have been developed in some American cities (e.g. Chicago, San Francisco, Portland, etc.) and widely in Canadian cities like Vancouver, Toronto or Montreal, it has not been implemented in hardly any European cities. In line with the implementation of buses with a high level of service (Heddebaut, 2010), in the last few years, some European cities have partially implemented it in some corridors, but not as a complete network. Some exceptions like Blue-buses from Stockholm or A-buses from Copenhagen are transit networks with a simple structure adapted to urban patterns. However, these represent the highest level of a wider hierarchical transit system, since the previous network was not removed completely. The resultant systems present a hierarchy since they are a mix between a transfer-based and a direct trip-based network. There was also a similar proposal for Barcelona that is currently being implemented (Estrada et al., 2011).

In this study, we want to justify under which circumstances transfer-based systems supply decentralized urban mobility scenarios better than direct trip-based network in terms of total system cost. At the same time, we want to provide guidelines to identify the best network structure for an idealized city with regard to its characteristics, transit system and demand. An analytical approach is used to compare different transit network structures for distinct degrees of urban decentralization. Four base structures are analyzed: radial and direct trip-based structures, and two transfer-based schemes: a grid and the hybrid structure presented in Daganzo (2010).

In the next section, our analytical model is presented for the case of a grid urban street layout. In Section 3, the transit networks are compared to understand their behaviors and to identify which is the best alternative in the different degrees of mobility decentralization. In Section 4, the effects on the applicability of each transit structure when main the input parameters of the model change are analyzed. Finally, our main conclusions are summarized in Section 5, and ideas for future research are included in Section 6.

2. Analytical model

To identify which type of transit network structure is the best alternative to serve demand, four base structures are compared. To make this comparison, we use a transit network design model with which the best configuration is defined for each network structure depending on different demand requirements. An analytical model is chosen for this task due to its compact formulation. Its simplifications of reality make the analysis of transit system behavior under different scenarios easier. In this section, the model used and its main assumptions are presented. This model was introduced by Daganzo (2010), and later extended by Estrada et al. (2011), Badia et al. (2014) and Chen et al. (2015).

2.1. City and mobility pattern

As with all analytical models, some simplifications of reality with regard to the extension of the territory that has to be served and the urban demand pattern are assumed. The city is considered as a square of side D , and its street layout is a grid. The average hourly demand of the transit system is λ and the demand at rush hour is Λ . These values are fixed: they do not vary with the level of service or other factors related to the transit supply. To represent the demand pattern, a trade-off between the simplicity of the analytical models versus real mobility patterns is needed. Most analytical models for the transit network design problem assume a uniform and independent distribution of generated and attracted trips over the whole area of study (Vaughan, 1986; Aldaihani et al., 2004; Daganzo, 2010; Nourbakhsh and Ouyang, 2012). However, this alternative makes it impossible to analyze different degrees of demand decentralization.

On the other hand, Vaughan (1987) considered more complex demand curves where the generation and attraction curves of demand were non-uniform and non-independent. This more realistic representation implies high complexity, and as a consequence, this type of models would lose its main benefit: compactness and simplicity. Badia et al. (2014) analyzed a ring-radial network, which only broke the simplest scenario with a non-uniformity in the radial direction; but this change generated complex formulations, and the results on the network configuration were not significant so as to obtain general insights.

Therefore, in order to conserve a simple and clear formulation at the same time that different scenarios of urban decentralization are represented, we divide the city into two areas. These are a central square of side ϕD , where $\phi \in (0, 1]$, and the remaining peripheral area between the boundary of that central area and the city's boundary. The generated demand is uniformly distributed over both areas but with different densities. The central density is greater than the peripheral density by a factor f_d . If this factor is equal to one, the demand is uniformly distributed over the whole city. If f_d is higher, the central area has a greater capacity to generate demand. However, only the central area is attractant, where the attracted demand is uniformly distributed. In addition, we maintain the independence between both demand curves. In this way, two types of trips exist with regard to the origin's location: (i) those whose origins belong to the central area, with a probability ρ , and (ii) those whose origins are peripheral, with a complementary probability $(1 - \rho)$. The value of ρ depends on the central area's size and the previous generated demand densities factor between both areas; that is, $\rho = f_d \phi^2 / [1 + \phi^2 (f_d - 1)]$.

In this way, these hypotheses allow us to determine the degree of demand dispersion easily by means of two factors. On the one hand, the size of the attractant central area represents in a simple way the different phases of the aforementioned urban form evolution. A small central area (i.e. low values of ϕ) is equivalent to a scenario where the CBD is the destination of all trips. A large central area, when ϕ tends to 1, represents a dispersed urban form where the CBD has lost its relevance with the appearance of new non-clustered centralities far away from it. Intermediate values of ϕ represent different demand decentralization degrees between these extremes, when the CBD starts to scatter over its surroundings. On the other hand, the different capacity of trip generation between central and peripheral areas allows the relevance of each area in the mobility to be distributed.

This demand distribution, where two different areas are distinguished, is similar to that of previous studies such as Tsekeris and Geroliminis (2013) and Li et al. (2013). The former analyzed the relationship between network structure and traffic congestion. The latter studied the density of radial roads related to residential distribution and housing prices. This monocentric approach takes into account different scenarios of CBD extension, but not a polycentric city where different centers are distinguished. This is an approximation where secondary poles appear and foment the occupancy of the territory between the CBD and itself, to become finally part of that extended CBD. Therefore, it is a monocentric approximation related to the prevalence of the traditional center in urban mobility described in the previous section.¹

2.2. Transit network structures

Figure 1 shows the four base structures studied: radial, direct trip-based, and two transfer-based systems, a grid and the hybrid scheme presented in Daganzo (2010). Decision variables of the radial scheme are the stop spacing s , which marks the route branching, and the headway H at the central point of the city for each direction (N-S and W-E). For the grid structure, the stop spacing s , which is coincident with the line spacing, and two headways are distinguished; H_c is the headway for lines that cross the central area and H_p is the headway for the remainder of the lines. This distinction guarantees a better deployment of resources. In the case of the hybrid network, three decision variables define it: the stop spacing s , the headway H for corridors at its central grid, and the parameter α that determines the size of

¹ A spatial econometric analysis of O/D matrices from some European cities shows that their demand patterns are characterized by a monocentric structure where only one cluster of high demand exists and is coincident with the traditional city center. This central area presents a demand density clearly higher than that of the rest of the city. In addition, the attracted demand has a higher concentration than the generated demand.

that grid. This final variable is the ratio between the side of the central gridded square and the side of the city. Therefore, radial and grid are the hybrid network for $\alpha=0$ and $\alpha=1$ respectively. These schemes are based on Daganzo (2010), where a more exhaustive explanation can be found.

The direct trip-based scheme deserves further explanation. Its main goal is to connect the attractant central area directly (without transfers) with the whole city. This system is an evolution of the radial network, where new lines are created to serve new attractant spaces resulting from the CBD's scattering. The resultant structure is like a superposition of different radial networks, each centered in one of the different attractant poles. The final stage of this structure is a ubiquitous system where every zone of the city is connected directly to every zone at least by one line; this happens when the attracted demand is distributed over the whole city ($\phi=1$).

For the idealization of this direct trip-based structure, we need to define a network scheme that satisfies this main goal. First, we have to define the demand poles, dividing the continuous attractant space into subareas. Then, we define the lines that connect each to the rest of city. The routing of these lines should avoid reiterative connections and try to cover the maximum area of service without tangled paths or loops. Therefore, the network is composed of different groups of lines, each connecting one subarea to the whole city. The lines of the same group are gathered in the subarea that they connect, and are dispersed out of this subarea to arrive at the rest of the territory. After considering different possibilities, we choose the structure presented in Figure 1b, since it accomplishes the previous requirements. This is explained in further detail below.

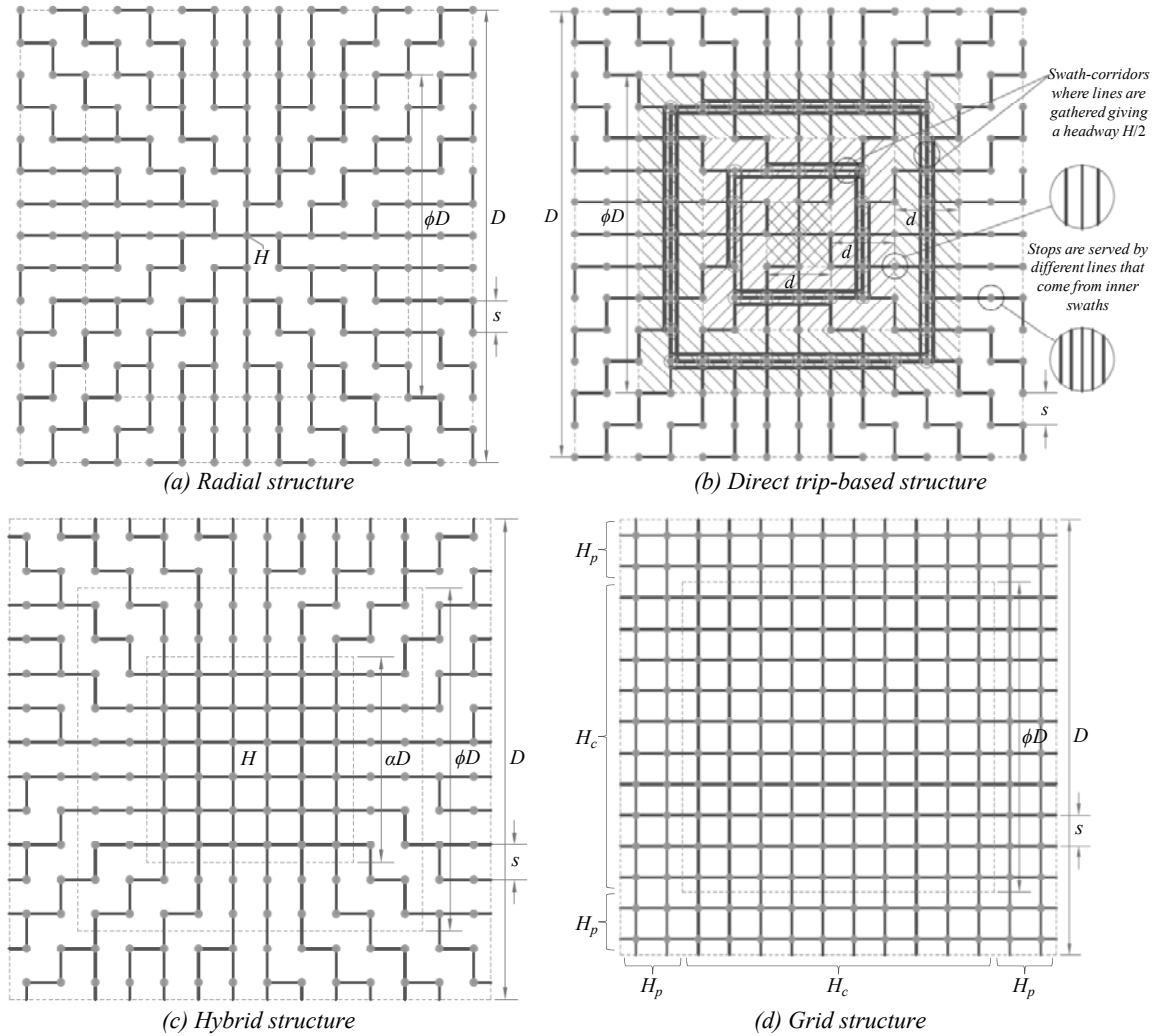


Figure 1. Different network structures compared.

The central attractant area is divided into different concentric swaths of width d . The central swath, a square of side d , is served by means of a radial network as in Figure 1a, whose lines are branched over the territory marked by the stop spacing s . The next swath is already connected to the central one by that radial network; however, new lines are needed to connect it to the rest of the city. To explain how the lines are arranged, we divide the swath into four zones, called swath-quadrants, and the area to be connected into four external-quadrants. These represent the area between the inner swath's boundary and the city boundary. Figure 2 shows these divisions, where we focus our attention on how the transit lines connect the East swath-quadrant.

Figure 2 shows three of the four groups of lines that serve the second swath and how they are arranged. Lines of the same group are gathered in one swath-quadrant and are distributed in the respective adjacent swath-quadrants for the next branching over the external-quadrants that they serve. This branching follows the same routing as the radial network of the central swath. The lines gathered in the East swath-quadrant (Figure 2a) are distributed over the North and South swath-quadrants and serve the respective external-quadrants. In this way, this group of lines completely connects the East swath-quadrant to these external-quadrants. To connect the West and East external-quadrants, we do not create new lines: rather, we use the lines gathered in the North and South swath-quadrants, as Figures 2b and 2c show respectively. Both groups of lines complement each other. Line L9 complements line L17; line L10 complements line L18; and successively until line L16 complements line L24. As a result, the East swath-quadrant is connected to the four external-quadrants.

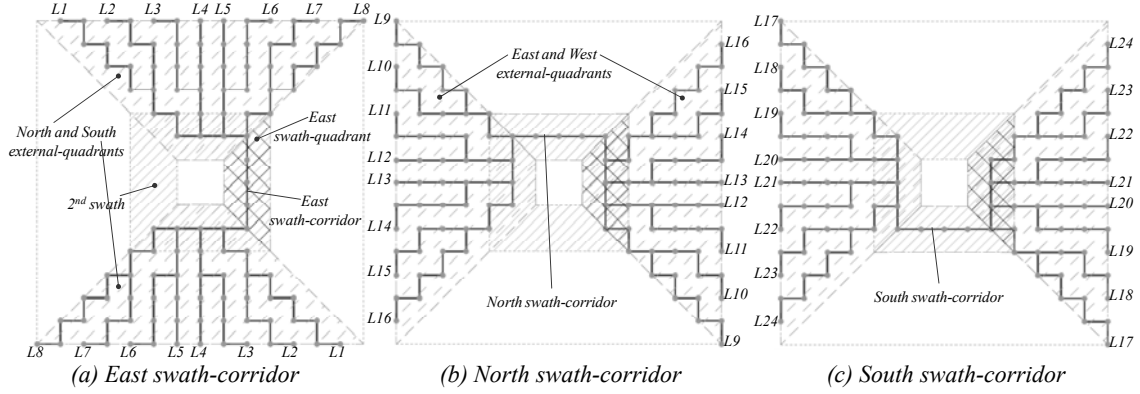


Figure 2. Detail of a direct trip-based network structure: how lines are arranged to connect the East swath-quadrant of the second swath to its external-quadrants by means of East, North and South swath-corridors.

By symmetry, rotating 90° , 180° and 270° each component of Figure 2, we obtain the lines that connect the other three swath-quadrants. For the remainder of swaths, the same explanation is valid. The lines are arranged in the same way, but they serve a smaller external area, since each swath is already connected to all its inner ones such as the swath explained before with regard to the central swath.

To summarize, the geometrical scheme of this structure is defined by two decision variables: stop spacing s and corridor spacing d , which is equivalent to swath width. In the external area of the city, which does not attract trips, transit corridors are not introduced due to their uselessness. On the other hand, the decision variable that defines the temporal coverage is H . This is the headway to the central point of the radial network, and, at the same time, is the resultant headway of those lines that belong to the same group; that is, the lines gathered in a swath-corridor of one swath-quadrant and serve its respective adjacent external-quadrants. This assumption guarantees the same level of service frequency for all the attractant area. Simultaneously, as Figure 2 shows, the swath-corridor is also partially crossed by lines that connect the same and the opposite external-quadrants. As a consequence, the resultant headway on these corridors is $H/2$.

2.3. Service operation: headways or schedules

A transit service can be operated by schedules or in headways. In the former, users know the moment when transit vehicles stop in their transit stops, and adapt their arrival based on that information. In the latter, users arrive at stop randomly due to the lack of that information. If trips are direct, a user only has to wait at an initial stop. In this case, he/she can adapt his/her arrival at that stop. However, if a transfer is

needed, a user has to wait at an intermediate stop where he/she cannot decide when to arrive. This is conditioned by the arrivals of the first and the second bus at the interchange point. Therefore, we only consider the possibility of working with both types of service operation in transit network structures that do not need transfers to complete the trips. Consequently, radial and transfer-based networks can only work in headways, but the direct trip-based structure can combine both service operations.

The approach followed to combine schedules and headways services is the same as that of Tirachini et al. (2010), for which three parameters are defined: (i) H_s is the cut-off headway between both types of service, which marks the boundary between inner and outer zones, working in headways and by schedules respectively; (ii) the safety waiting time h_s is a fixed value for all stops that works by schedules; the user arrives some minutes before the arrival time of the vehicle; and (iii) home waiting time as an opportunity cost, which is a variable time defined as the product of the headway of the service at the stop and the factor f_s . This factor has a value less than 0.5, since waiting at home/work is less negatively perceived than at a transit stop. Therefore, Function (1) determines the waiting time at bus stop w_s for a line with headway H_l .

$$w_s = \begin{cases} H_l/2, & \text{if } H_l < H_s \\ h_s + f_s H_l, & \text{if } H_l \geq H_s \end{cases} \quad (1)$$

In a direct trip-based network, another parameter ε_H is used to mark the square cordon of side $\varepsilon_H D$. This is the boundary between both types of operation. The system operates in headways in the central area inside that cordon, and by schedules outside. If ε_H is equal to 1, all the network works in headways. On the contrary, the service is mixed.

2.4. Model formulation

To determine the best network structure, we compare the total system cost of each. Therefore, the different components of this cost are estimated. From the agency's point of view, the model computes three costs: infrastructure length L (km), kilometers traveled V (veh-km/h) and fleet M (veh-h/h). Their derivations are basically defined by the network structure. On the other hand, access A , waiting W , in-vehicle T and transferring $\delta e_T/w$ (δ (km) is a transfer penalty such as length, e_T is the average number of transfers per trip and w (km/h) is the pedestrian speed) times are computed as user costs. These costs are the different components of the transit chain, which is the path that users follow in the network from their origins to their destinations. Some criteria are accepted to define this path, which is the first step in deriving these user costs.

- Access and egress cost: The user takes the bus at the closest stop to his/her origin and gets off at the closest stop to his/her destination, conditioned by the variable s . There is only one exception in a direct trip-based network: the initial or final stop is the closest stop located at one of the swath-corridors where the lines are gathered, conditioned by the variable d . In this network, we assume that the user does not make transfers.
- Waiting cost: This is conditioned by the previous assumption. However, the headway varies smoothly among nearby stops. One exception exists in the grid network, where stops in some peripheral areas are served by two perpendicular lines with different headways. In these stops, the users choose the path of the line with the minimum headway.
- In-vehicle cost: The user opts for the shortest path that the network structure allows between its origin and destination.
- Transfer cost: The user chooses the path with the minimum number of transfers.

In general, assuming the first two criteria, a path exists that accomplishes the third and fourth criteria, which has the shortest total travel time. Figure 3 shows the path for different origin-destination pairs for the four structures. In a radial network, the path that connects any origin-destination pair accomplishes the four criteria, either by direct trips if the origin and destination belong to the same line (O_1-D_1), or by one transfer at a central stop ($O_2-D_{2,1}$ or $O_2-D_{2,2}$). It is possible to observe that some trips are longer than the real distance due to the structure itself. The user must arrive at the center of the network to take the line that serves the destination; for example, the $O_2-D_{2,1}$ pair.

In a grid, the best path for any origin-destination pair is the shortest in a L_1 metric, and at the same time, it has the fewest transfers: zero or one. In the area served by central and peripheral lines, the users take the line that implies the shortest waiting time; by definition H_c is always lower than H_p . The trip O_8-D_8 is an example of this. As the hybrid network is a mixed scheme, it presents the same type of paths as the grid structure when the origin and destination are located in its central grid (O_5-D_5 and $O_5'-D_5'$), or as the radial network when one of them is in the peripheral branching (O_6-D_6 and $O_6'-D_6'$). However, with one exemption, some trips need a second transfer: O_7-D_7 is an example. In these trips, the origin and destination are located on branched sections and belong to the same or opposite quadrants. These trips only exist when the variable α is smaller than the parameter ϕ . This relationship determines the types of trips that exist.

In a direct trip-based network, the third and fourth goals are coincident when the destination is located in the same quadrant as the origin or in adjacent quadrants, for example O_3-D_3 and $O_{4,1}-D_4$ trips. However, this does not happen when both extremes of the trip belong to opposite quadrants ($O_{4,2}-D_4$). In this case, some direct trips that go through the destination swath are longer than alternative paths that cross the city center where the users must make one transfer. The best alternative depends on different aspects like commercial speed, transfer penalty, user perception of in-vehicle and transfer times, and in which swath the destination is. For simplicity, we consider that the users choose the direct trip although this is longer, since the main objective of this network is guaranteeing direct connections between origins and destinations.

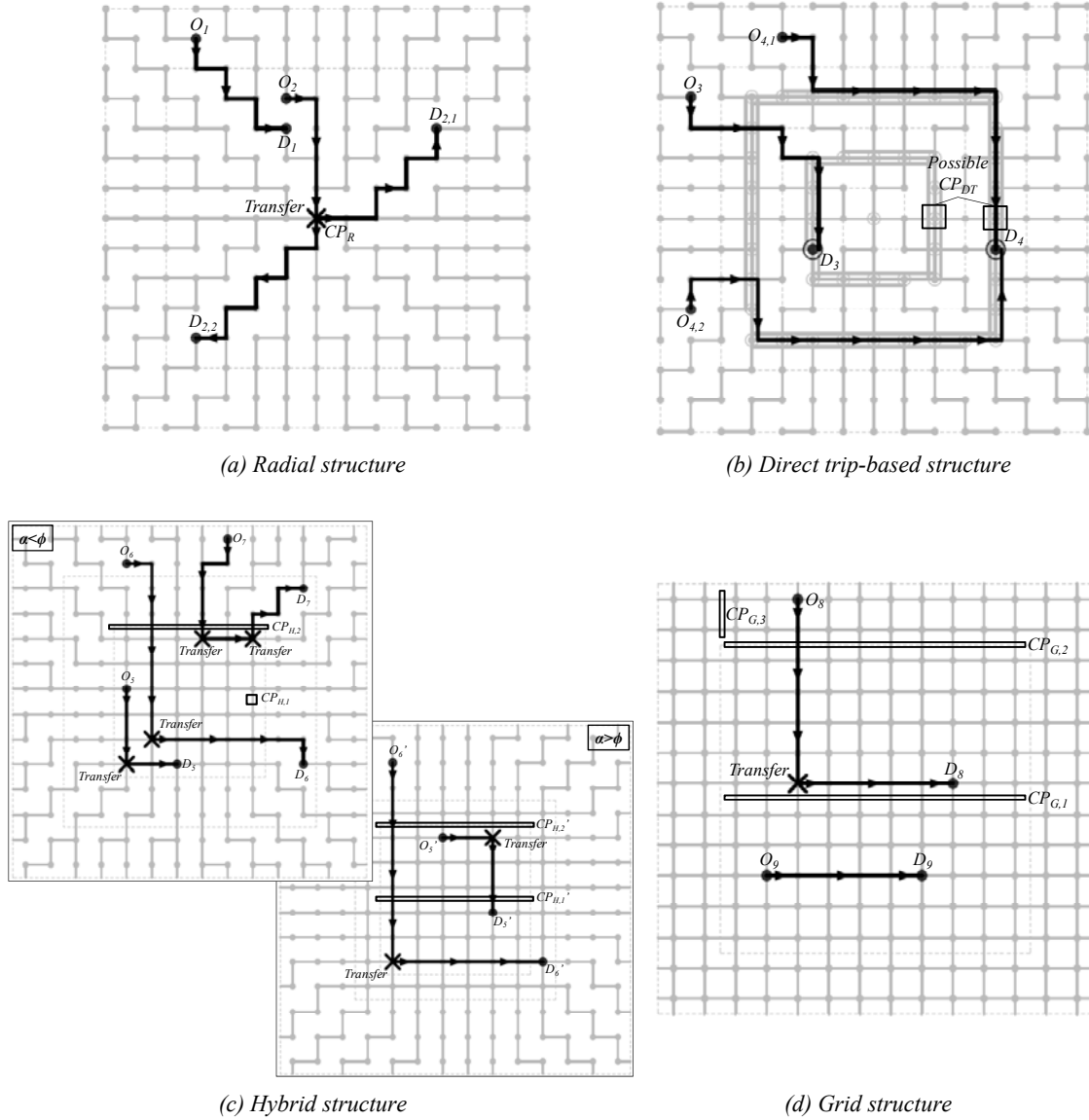


Figure 3. Paths and critical loaded points of the structures studied.

The formulations are derived in detail in Appendix A as a function of the corresponding decision variables for each analyzed structure. Each component of the system cost is included in the objective function (2), which is minimized to find the optimal system configuration. In the objective function, all the costs have to be expressed in the same units; in this case, hours per passenger. For this reason, agency costs are multiplied by their respective unit costs ϵ_i and divided by the value of time μ and the average hourly demand λ . On the other hand, the time perception for different costs of the user transit chain is taken into account by means of weights w_i . Finally, constraints (2a) related to the decision variables and the vehicle capacity condition the optimal solution.

$$\min\{Z = C_A + C_U = [\epsilon_L L + \epsilon_V V + \epsilon_M M]/\lambda\mu + [w_A A + w_W W + w_T T + w_t(\delta/w)e_T]\} \quad (2)$$

$$\begin{aligned} \text{s.t. } & s > 0; H > 0; s/D \leq \phi; 0 \leq C, \text{ for radial network} \\ & d \geq s > 0; H > 0; s/D \leq \phi/2; d/D \leq \phi/2; 0 \leq C, \text{ for direct trip-based network} \\ & s > 0; H_p \geq H_c > 0; s/D \leq \phi; 0 \leq C, \text{ for grid network} \\ & s > 0; H > 0; s/D \leq \min\{\alpha; \phi\}; 0 \leq C, \text{ for hybrid network} \end{aligned} \quad (2a)$$

Regarding the constraints, all decision variables have to be positive for physical reasons, and the maximum vehicle occupancy O must be lower than the vehicle capacity C . Moreover, a ratio between stop or corridor spacing and the length of the side of the central attractant area, or the central grid for the hybrid network, must be respected as a consequence of the derivation of the formulation.

To estimate the occupancy, the most loaded points in each network are compared. Figure 3 shows the location of these points in the different networks. These are: the central point in a radial scheme (CP_R); and three points in a grid structure: the middle point ($CP_{G,1}$), the point at the attractant central area boundary ($CP_{G,2}$) of one central corridor, and the point $CP_{G,3}$ of the peripheral corridors. In the hybrid network, it depends on the ratio between α and ϕ . If α is greater, the critical points are located at the boundary ϕD and at the central point of the corridors that cross the attractant area: points $CP_{H,2}'$ and $CP_{H,1}'$ respectively. On the contrary, at the boundary αD ($CP_{H,2}$) and at the central point of the most external corridors ($CP_{H,1}$). Finally, for a direct trip-based network, the critical point belongs to one swath that can be the most external one when the attractant area is small, or an intermediate swath when this area is larger. Appendix A contains a more detailed explanation and the respective formulae.

The objective function for all network structures is convex. This characteristic makes its minimization trivial; the local optimum found is the global solution. Its optimization is performed by means of a grid search. This simple method is applicable due to the small number of decision variables of the model. The maximum number of variables in one network structure is three. For example, in the direct trip-based structure, a computer Intel Core™2 Quad CPU Q9550 @2.83 GHz obtains the optimal solution for a specific case of study in 15 min, where a long range of values with a small pace is attempted for the different decision variables: $H \in [0.05, 5.0]$ min and $\Delta H = 0.05$ min; $s \in [0.005, 1.0]$ km and $\Delta s = 0.005$ km; $d \in [0.005, 2.0]$ km and $\Delta d = 0.005$ km. These ranges cover a large number of values that are found in real networks, and the steps are sufficiently fine with regard to the real margins that geometrical design or service operation allow.

3. Case study

In this section, a non-specified city is analyzed to show the behavior of the different structures with regard to demand decentralization. Section 3.1 demonstrates this behavior in detail considering a constant generated demand density over the whole city. This assumption allows that the decentralization degree is only defined with ϕ , which ranges from low values around 0.1–1. In Section 3.2, we distinguish the generated demand density between the central area and the periphery, showing how a higher concentration of origins in that central area changes the results of the model.

Table 1 summarizes the model's input parameters for this case study. These are the same as the hybrid model (Daganzo, 2010; Estrada et al., 2011; Badia et al., 2014) with the introduction of the parameters related to the service operation of Section 2.3. In addition, we consider two types of vehicles: standard buses and articulated buses. The former is cheaper, but the latter is suitable to serve network configurations or urban centralization scenarios with a high passenger load at the critical points of the network. In any case, the fleet is composed only of one type of vehicle; all standard or all articulated.

Table 1. Characteristics of the case study.

Input parameter	Variable	Units	Value
Demand at rush hour	λ	p/rh	50,000
Average hourly demand	λ	p/h	20,000
Square city dimension	D	km	7
Value of time	μ	€/h	15
Equivalent penalty distance per transfer	δ	km	0.3
Free flow speed	v_{FFS}	km/h	30
Walking speed	w	km/h	4.5
Unit infrastructure cost	ϵ_L	€/km-h	76
Unit distance cost	ϵ_V	€/veh-km	0.85 or 1.1 ^a
Unit vehicle cost	ϵ_M	€/veh-h	35 or 36 ^a
Vehicle capacity	C	p/veh	80 or 150 ^a
Dwell time	τ	s	35
Boarding (and alighting) time	τ'	s	3
Time perception weight of access	w_A	-	2.2
Time perception weight of waiting	w_W	-	2.1
Time perception weight of traveling	w_T	-	1.0
Time perception weight of transferring	w_i	-	2.5
Cut-off headway between types of service	H_s	min	12
Safety waiting time	h_s	min	5
Home waiting time factor	f_s	-	1/12

^a The first value corresponds to a standard bus and the second to an articulated bus.

3.1. Constant generated demand density over the whole city

As the generated demand density is constant over the whole city, the factor between the central and the peripheral densities is $f_d = 1$. Therefore, the weight of the trips whose origin is central is $\rho = \phi^2$. Figure 4 shows the total, agency and user costs with regard to the demand decentralization degree ϕ for the different transit network structures: radial (*Ra*), direct trip-based network (*DT*), grid (*Gr*) and hybrid (*Hy*).

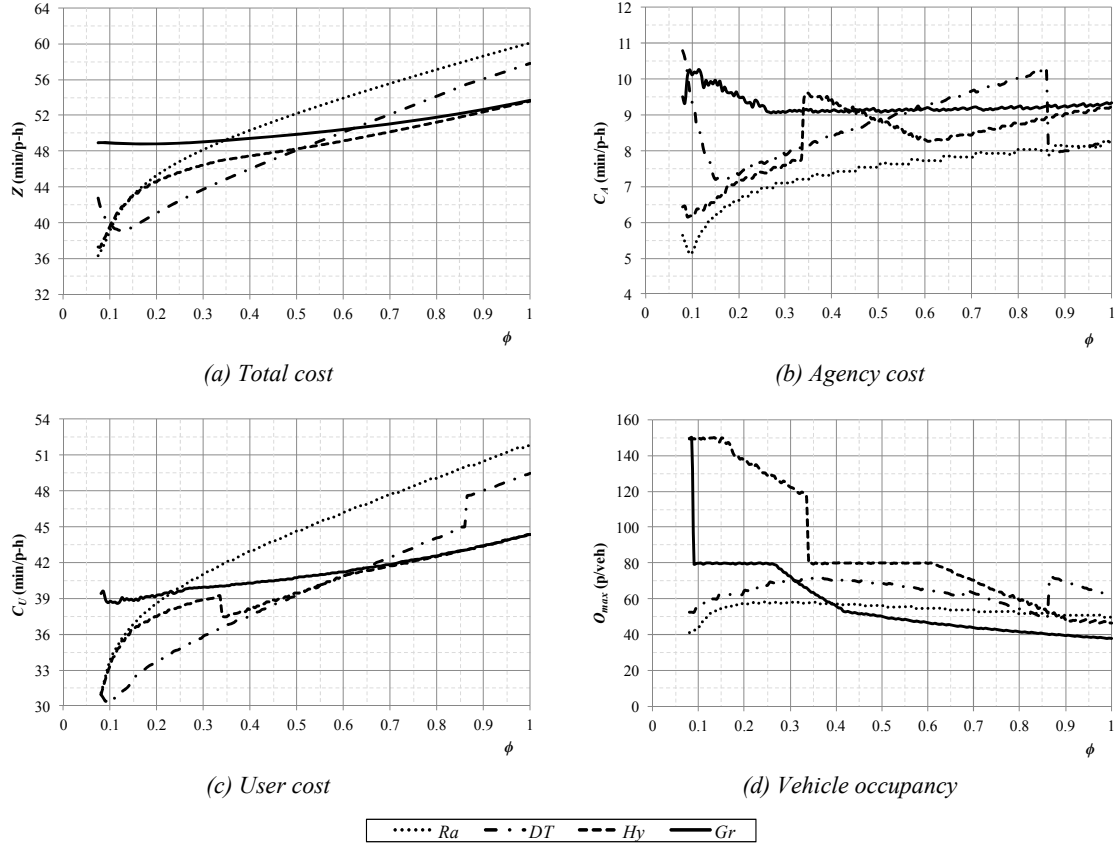


Figure 4. Evolution of costs and vehicle occupancy with regard to the demand decentralization degree parameter ϕ .

As observed in Figure 4a, the best transit structure is not always the same: it depends on mobility requirements. There are different cut-off points between the total cost curves of the different transit

network structures. These points mark the ranges of applicability of each structure; that is, which is the optimal alternative for each stage of the demand dispersion process.

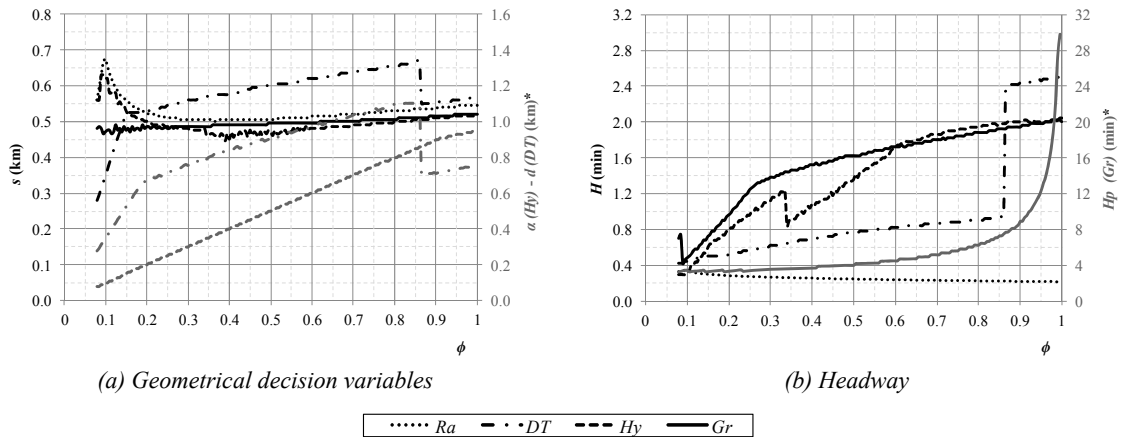
A radial network is the most suitable alternative when the vector of displacements is centripetal, i.e., people want to travel to a small central area. However, its total cost increases sharply when the demand starts to scatter. Therefore, the transit system needs a change of its network structure. Two options are possible: either a direct trip-based structure by means of the introduction of new lines over the radial network, or a hybrid scheme with an initial transformation of the central part of the network to a grid. As Figure 4a shows, the former is the better solution, since it keeps the total costs lower compared to the latter. This strategy supposes a greater agency investment, which is compensated by the reduction of the increasing user cost that demand decentralization implies.

From that point on, the direct trip-based structure continues to remain the best alternative until the demand decentralization reaches a higher degree. In our case study, this value is 0.52. If that decentralization degree is exceeded, a change from a direct trips strategy to a hybrid network is the best decision. This change guarantees that the transit network works in the most efficient way. The hybrid network is the best for higher levels of decentralization.

All structures have increasing cost with demand dispersion. These curves vary smoothly. However, the evolution of the agency and user costs (Figures 4b and 4c respectively) shows break points and changing tendencies along the decentralization process. Three factors explain these changes: geometrical decision variable constraints, the vehicle capacity constraint, and which type of service operation is used. For low values of ϕ around 0.1, the stop/line spacings are restricted to short distances, as Figure 5a shows. Its main consequence is a sharp growth in agency cost.

The second factor produces the jumps at $\phi = 0.09$ and 0.34 in the grid and the hybrid network respectively (Figures 4b and 4c). For lower values, the high pressure on the small attractant area forces these networks to work with articulated vehicles. However, when the attracted demand is spread over a larger area, the occupancy at the most loaded points decreases, and a change of vehicle size is possible. As Jara-Díaz (2003) and Jansson (1980) have shown, if the objective function includes the user cost, working with smaller vehicles is better due to a better trade-off between agency and user costs. The former increases due to a larger fleet, and the latter reduces because of lower headways (Figure 5b). In both networks, from these decentralization degrees, they work at capacity, reducing progressively the agency costs until the vehicle occupancy is no longer a constraint: $\phi = 0.27$ and 0.61, respectively.

The final factor only affects the direct trip-based structure. At a decentralization degree of around 0.86, the agency and user cost curves change sharply. From this stage, this structure starts to work by schedules in most of its stops. Therefore, almost all trips then work by schedules.



* Gray curves are related to the vertical axes on the right, and their line style to the network structure of the respective decision variable.

Figure 5. Evolution of decision variables with respect to the demand decentralization degree parameter ϕ .

Figure 5 shows the evolution of the decision variables for all network structures. Their curves are also conditioned by the previous factors; hence, some abrupt changes are manifest. Figure 6 includes two metrics of the network performance from the user point of view: average number of transfers per trip, and average in-vehicle travel time per trip. This information allows us to understand the main weaknesses and advantages of each network. The most important weakness of the radial structure is the length of its trips, as these pivot around the city center. This increases in a constant way, and is clearly greater than the other structures. Moreover, the number of transfers increases rapidly at the initial steps of the demand decentralization, until it reaches similar values to transfer-based networks.

The main inconvenience of a direct trip-based network is that, to avoid transfers, the network configuration tends to longer stop and corridor spacings and higher headways, as Figure 5 shows. Therefore, users progressively lose spatial and temporal coverage to guarantee direct connections. Otherwise, a larger volume of resources would be needed to serve the attractant areas directly with the whole city when the demand dispersion grows. However, there is a dispersion degree from which this tendency is no longer sustainable. As a consequence, the system starts to work with mixed services, and the network configuration changes to closer stops at the expense of service frequencies.

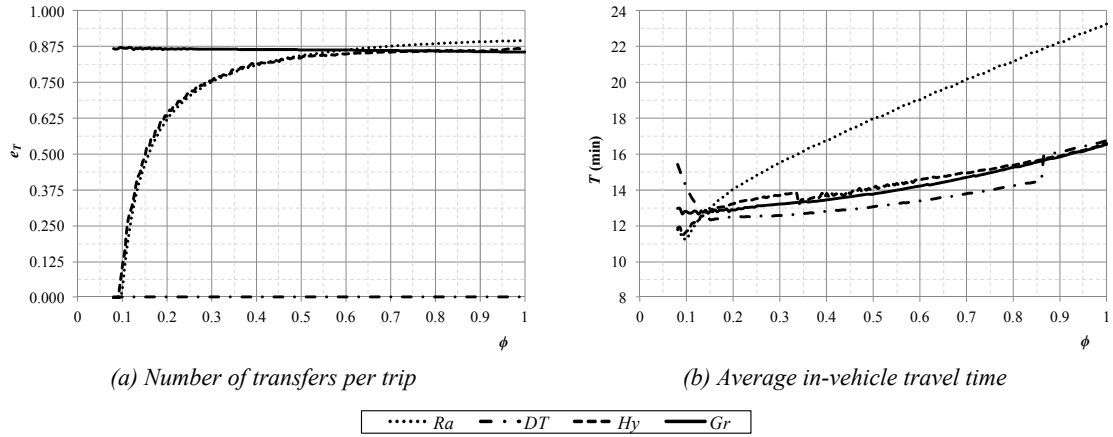


Figure 6. Evolution of the number of transfers per trip and the average in-vehicle travel time with respect to the demand decentralization degree parameter ϕ .

A transfer-based network structure has constant spatial coverage; however, the headway of service increases, since there are more corridors where the fleet is allocated. Its main weakness is that the transfer is an essential step of the transit chain. Between the hybrid and the grid structures, the former is always a better alternative. This is focused on its central grid, which evolves with the size of the central attractant area. Figure 5a shows it with increasing values of α . This fact allows the number of transfers to be kept low, and better deployment of resources. The number of transfers in a grid is practically constant regardless of the mobility pattern, and different headways for central and peripheral corridors are insufficient to compensate the double coverage for all stops, one horizontal and one vertical corridor.

3.2. Different generated demand densities between central and peripheral areas

We now let the value of f_d change from 1, as in the previous section, to 30, a scenario where the central area has a higher trip generation capacity than the rest of the city. Figure 7a shows that the total system cost decreases when the central area is denser. However, as Figure 7b shows, the area of applicability among the different structures is practically constant, since the decreasing tendency of the total cost is similar in all structures. The value of ϕ , where a change of structure is justified, varies between 3.4–7.1%, from $f_d = 1$ to $f_d = 30$, between the radial and the direct trip-based networks and between this second and the hybrid scheme, respectively.

Therefore, regarding the demand representation, the results show that the size of the central area is the most important determinant of the applicability of each structure. That is, the area where we have to develop swath-corridors or a central grid. However, the fact that greater or fewer users come from the periphery or travel within the central area is not as relevant. The introduction of more complex demand representation, which now depends on a second parameter, does not provide important insights.

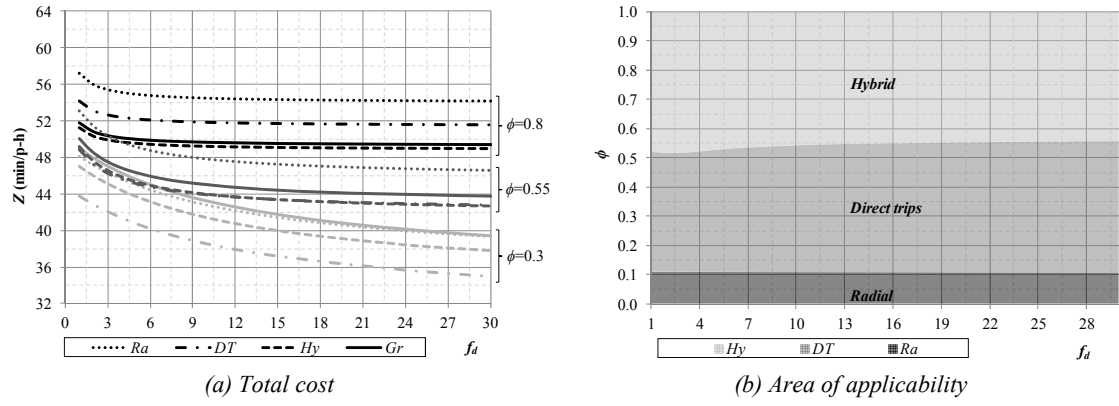


Figure 7. Total cost and area of applicability with regard to the factor between central and peripheral generated demand densities f_d .

4. Variations in the demand decentralization degree of change of structure due to input parameter variations

The previous section showed that the best transit network structure is not always the same: it depends on how the demand is spread over the city, i.e., whether it is concentrated or dispersed. Therefore, there exists a demand decentralization degree that justifies a change to the transit network structure. We now analyze how the cut-off point of change among the different structures varies with regard to four main characteristics: level of transit demand, city size, transfer penalty and ratio between agency and user unit costs. We use the previous base scenario of Section 3.1, where we only work with ϕ to determine the urban decentralization degree. For the case of the level of transit demand, the ratio Λ/λ is always considered equal to 2.5. On the other hand, the city size is represented by the square dimension D .

The structures analyzed here are: radial (Ra), direct trip-based (DT) or hybrid (Hy). The grid structure is removed in this part of the analysis since it is never the best solution, which was shown in the previous section. The results are presented in two different figures. Figure 8 shows the values of ϕ of change between different network structures. Not all the boundaries between structures are included on the figures; only those that have relevance, since they delimit structures that interchange the position of minimum cost. On the other hand, Figure 9 presents the area of applicability of each structure when each of the four input parameters analyzed is examined independently.

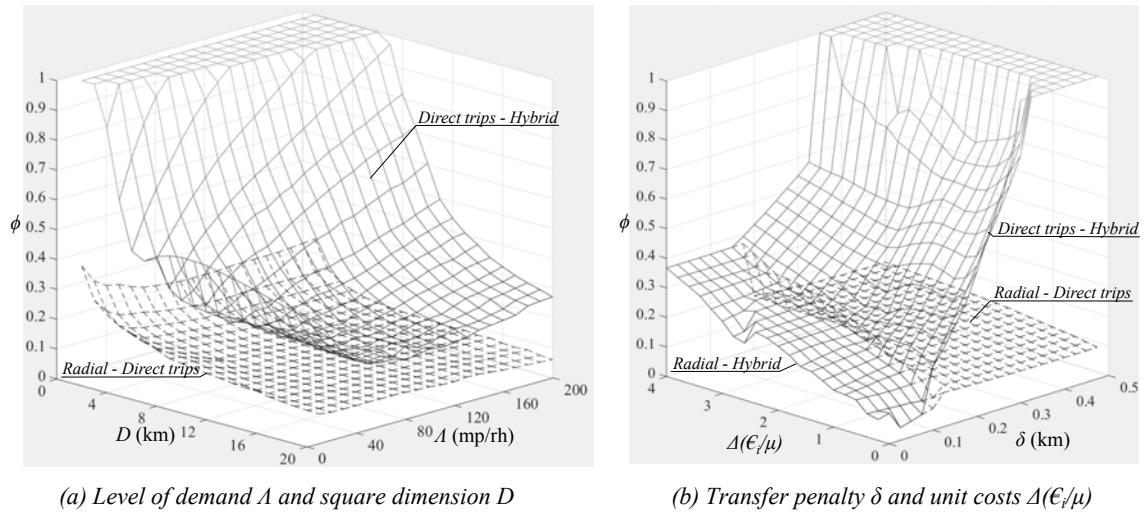


Figure 8. Evolution of the demand decentralization degree ϕ as a cut-off point between transit network structures with respect to the four input parameters analyzed.

The results of Section 3 are reinforced here. The radial structure is the most suitable solution for high levels of demand concentration. When this alternative is overtaken, the solution lies in introducing new lines to conform to a direct trip-based structure. The boundary between radial and hybrid networks is not

relevant unless the transfer penalty is low, as Figures 8b and 9d show. In this case, the purpose of a direct trip-based structure loses its meaning. Subsequently, when the demand decentralization reaches higher values, transit systems have to face a reorganization of their network structures to implement a transfer-based structure.

Figure 8a shows how the areas of applicability change when the transit demand and city size vary. Regarding the level of demand, the direct trip-based structure increases its area of applicability when demand grows. First, this network has a lower ratio of C_U/C_A than the radial network. Secondly, against the hybrid structure, increasing demand justifies the development of more independent lines, maintaining a suitable spatial and temporal coverage to connect all trips directly for higher demand decentralization degrees. However, this tendency changes for low levels of demand. The direct trip-based structure gains applicability versus the hybrid network, since the former operates by schedules in these scenarios – a type of service that reduces the agency investment at the expense of the users.

On the other hand, radial and the direct trip-based structures reduce their applicability when the city is larger. The former is penalized because it highlights one of its weaknesses, the trip length. The latter would need a large number of lines to supply a good level of service, and as a consequence, a huge volume of resources. On the contrary, the synergies among transit lines in a transfer-based structure allow better performance to be supplied with limited increases in investment.

The higher the value of one of these parameters, the smaller the effects that the other produces when it changes. This phenomenon is more evident for the demand variations conditioned by the city size. For small cities, demand growth increases the suitability of direct trips. However, this fact happens in a moderate way when the city is larger. In this instance, the direct trip-based structure gains applicability for low levels of demand.

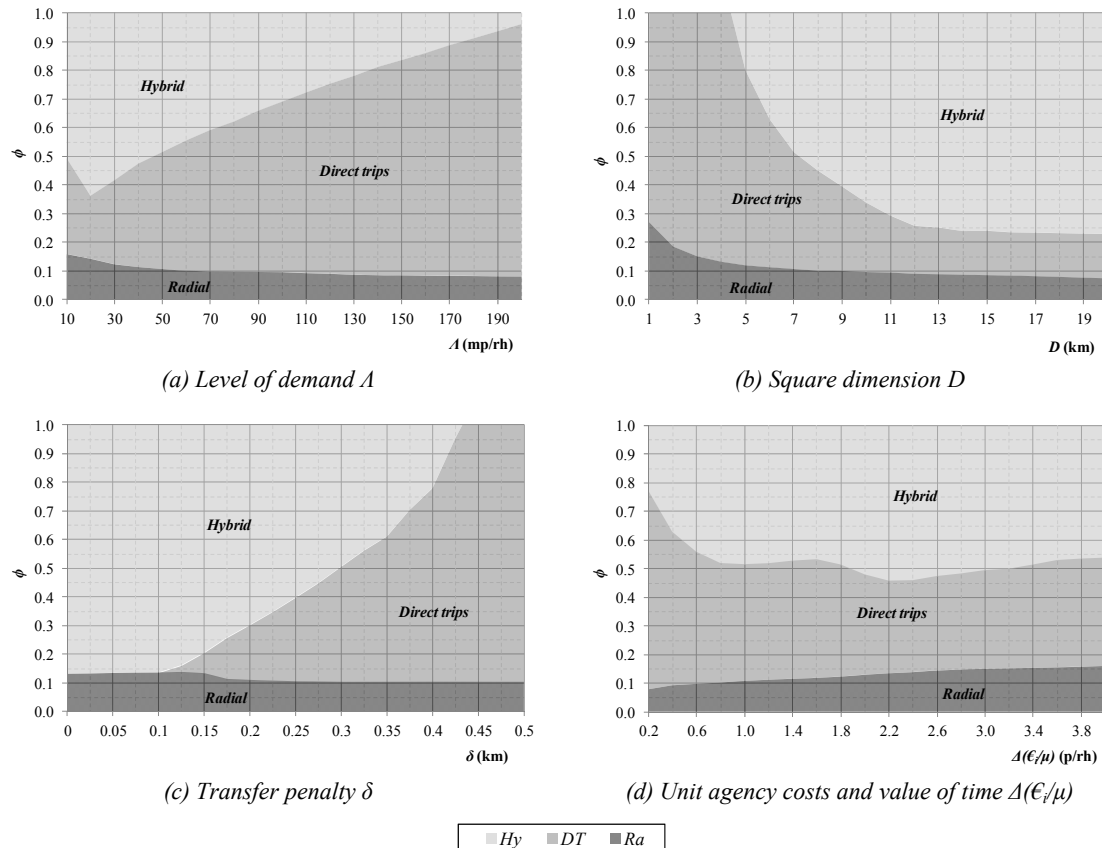


Figure 9. Area of applicability with regard to the four input parameters analyzed for the base scenario of Section 3.1.

Figure 8b focuses attention on the transfer penalty, to evaluate the interchange stop design, and the ratio between unit agency costs and the value of time, i.e., ϵ_i/μ where i is L , V or M . These ratios vary from the base scenario by multiplying them by a factor of 0.2–4. In this way, different scenarios are considered:

those where the value of time prevails over the unit costs of infrastructure, kilometers traveled and fleet, and those where the value of time is undervalued versus the unit agency costs.

As observed, the former has a greater impact on the cut-off point between the direct trip-based structure and the hybrid scheme. There exists a range of transfer penalties, which approximately varies from 0.1–0.45 km, where this cut-off point varies sharply. Beyond this range, one of these structures is never the best solution: direct trips when the transfer cost is lower, or a hybrid network when the penalty is higher. Changes due to unit costs are moderate. Figure 9d shows that for scenarios where the user cost prevails, a greater agency investment is justified; therefore, direct connections are supplied without losing other performances of the transit system. The network completely operates in headways for higher levels of dispersion.

Regarding the radial structure, its applicability is practically constant. The most significant change takes place when the agency costs gain relevance against the user costs, especially when this structure competes with the hybrid scheme. The reason is the low agency investment that characterizes the radial network.

5. Conclusions

The analysis developed throughout this study has confirmed that the best transit network structure is not always the same: it depends on the mobility pattern of the city. We have identified the relationship between the demand requirements and the applicability of different transit network structures. A comparison among four base network structures identifies three different scenarios of demand decentralization. Each is related to one of these structures: (i) High demand concentration in the central district, where a radial system is the most suitable solution; (ii) an intermediate scenario when the demand starts to scatter around that central district; the best alternative is the development of new lines to connect directly the new attractant areas; and finally (iii) a dispersed demand pattern, where the hybrid scheme as a transfer-based structure is the most efficient alternative.

The reasons that justify these changes of network structure are the strengths and weaknesses that characterize each of the previous three structures. The radial network is cheap from the agency's point of view, but the number of transfers and the trip length increase sharply with the demand decentralization. A service based on direct connections removes the transfers completely and limits further increases of trip length. However, keeping those direct trips in a decentralization process implies great investment, at the same time as negatively affecting the spatial and temporal coverage. In dispersed scenarios, the hybrid network allows for better utilization of the resources by means of transfers, which at the same time is its main disadvantage.

On the other hand, the size of the central area is the most important factor that determines the applicability of each structure. There exists a size of this area from which the development of the direct trip-based network is not the most efficient measure. However, a change of structure to a hybrid scheme is the best decision. Therefore, ϕ is more relevant than f_d in representing the demand decentralization degree.

The other aspect from our results is that the point of change among the different transit network structures is not constant, and all three aforementioned phases do not always exist. This depends on the characteristics of the city, the transit system and the demand. The most important variations are produced by the city size and the transfer penalty. Large cities and well-designed transfers favor a transfer-based network for lower degrees of dispersion. However, in some scenarios, the level of demand also plays a role. High levels of demand justify greater investment in direct services for higher scenarios of decentralization; on the contrary, the transit system works by schedules when the number of users is low.

These results agree with previous discussions about transit network structure. As Newell (1979) noted, when the attractant areas of a city are so concentrated around a central space, the best solution is a radial (or hub and spoke) scheme. However, at the same time, our comparison shows that a grid becomes a better option for scenarios where peripheral areas have more prominence on mobility; this was already advocated by Thompson (1977). In addition, the intermediate hybrid network is in line with the previous arguments of Newell (1979) and Thompson (1977). Transit lines are focused on a central area where the activities are concentrated, as Newell (1979) argued. However, that central area is no longer a small

center or corridor: it becomes ever larger with the evolution of the city. Therefore, it is convenient that this central area has to be served by a grid, as Thompson (1977) proposed.

In conclusion, by means of the analytical comparison proposed here, we have obtained results that allow us to determine guidelines to make decisions about how transit systems must be designed. We have identified the best transit network structure among different alternatives, and the suitable moment when the network configuration has to change to increase the efficiency of the transit system. However, some aspects, such as other street layouts or hierarchy of transit services, were not considered; these factors could affect our results. There is therefore scope for future extensions of this work.

6. Future works

From this point, two lines of research are possible. In the first place, we could continue with the analysis of the transit network design with more complex systems by means of hierarchical and mixed structures. In addition, we could analyze these structures on different street layouts to understand how the street layout affects the results of the study.

On the other hand, the second line concerns how we could transfer the analytical results to real cities. The question that has to be answered in the future is whether cities, or which cities, have the characteristics that justify a reconfiguration of their transit networks from radial or direct-trips based structures to transfer-based ones. That is, if the evolution of these cities has changed their urban mobility patterns to more dispersed scenarios. Our analytical approach would be used as a tool to evaluate whether a change of structure is a good decision to improve the competitiveness of transit networks.

Acknowledgments

H. Badia was funded by an FPU grant, reference number AP2010-4544, from the Spanish Ministry of Education.

Appendix A. Model's formulation

Nomenclature	
H	Headway [h]
s	Stop spacing [km]
d	Swath width or swath-corridor spacing in a direct trip-based network structure [km]
α	Parameter defining the central grid in a hybrid network structure
Z	Total system cost [h/p-h]
C_A	Agency cost [h/p-h]
C_U	User cost [h/p-h]
L	Infrastructure length [km]
V	Kilometers traveled per vehicle and hour [veh-km/h]
M	Number of vehicles working per hour [veh-h/h]
A	Average access time per trip [h]
W	Average waiting time per trip [h]
T	Average in-vehicle time per trip [h]
e_T	Average number of transfers per trip
O	Vehicle occupancy in the most loaded points of the network [p/veh]
ϕ	Demand decentralization degree
f_d	Factor of densities between central and peripheral areas
ρ	Portion of generated demand at central area
λ	Demand during the rush hour [p/h]

λ	Average hourly demand [p/h]
D	Length of the side of the square that represents the city [km]
μ	Value of time [€/h]
δ	Equivalent penalty distance per transfer [km]
v_{FFS}	Speed without considering spent time at stops [km/h]
w	Pedestrian speed [km/h]
ϵ_i	Unit agency cost, where $i = L, V$ or M [€/km-h; €/veh-km; €/veh-h]
C	Vehicle capacity [p/veh]
τ	Dwell time per stop [h]
τ'	Boarding (and alighting) time per passenger [h]
w_i	Time perception weight, where $i = A, W, T$ or t
H_s	Cut-off headway between types of service [h]
h_s	Safety waiting time [h]
f_s	Home waiting time factor

The model's formulation for the four structures is presented in detail here. First, radial, grid and hybrid formulations are directly introduced in Section A.1. Their derivations are very similar to those presented for the hybrid model in Daganzo (2010) and Estrada et al. (2011). For this reason, a short explanation introduces these formulations. This appendix includes some indications, but excludes exhaustive derivations and proofs. We direct readers to Daganzo (2010) and Estrada et al. (2011) for more detailed demonstrations of these formulae. Here, the formulae depend on two new parameters, the central decentralization degree ϕ and the percentage of generated demand at central area ρ . This latter parameter is defined as $\rho = f_d \phi^2 / [1 + \phi^2 (f_d - 1)]$. Next, the direct trip-based network's formulation is derived in depth in Section A.2. Figures 1–3 help to understand the explanations and proofs. The network structure and the paths followed by the users are essential information for the derivations.

A.1. Formulation of the radial, grid and hybrid network structures

Tables A.1 and A.2 summarize agency and user partial costs for the radial and grid structures. Regarding agency costs, their derivations follow the procedure of Daganzo (2010). For the radial scheme, the formulations are derived such as the peripheral area of the hybrid network. For the grid, the derivation procedure is equal to the central area of the hybrid model. However, one modification is considered in this case: the network works with two different headways, one central H_c and other peripheral H_p .

Table A.1. Formulation of partial costs, commercial speed and occupancy constraint for radial network.

Partial cost	Radial network
L	$L = D^2 / s$
V	$V = 6D / H$
M	$M = V / v_c$
v_c	$v_c = 1 / [1 / v_{FFS} + \tau / s + \tau' (1 + e_T) \Lambda / V]$
A	$A = s / w$
W	$W = H [15D(1 + 2\phi + 2\phi^2) - 15s(1 + \phi) - \rho(15D + s(1 + \phi))] / 45s(1 + \phi)$
E	$E = [15D(1 + 2\phi + 2\phi^2) - 15s(1 + \phi) - \rho(15D + s(1 + \phi))] / 30(1 + \phi)$
T	$T = E / v_c$
e_T	$e_T = 1 - s(3 + \rho) / 3\phi D$
O	$O = (SF)H\Lambda[6\phi D - s(3 + \phi^2)] / 24\phi D$

The user costs have two main modifications from Daganzo (2010): (i) the first is caused by the demand distribution, and (ii) the second is a consequence of the estimation of the average number of transfers per trip e_T . Regarding the former, the demand distribution changes the estimation of user costs because it changes the O-D matrix. The new parameters ϕ , f_d and ρ are introduced in the formulation to control this

phenomenon. Two zones are distinguished: a central area, the origin of a percentage of trips and destination of every trip; and a peripheral one, the origin for the remainder trips. Therefore, only two types of trips exist: (i) those whose origin and destination are located at the central attractant area, and (ii) those whose origin is peripheral but their destination is central. In the case of the number of transfers, the same assumption as made by Estrada et al. (2011) is accepted to estimate e_T : Trips whose origin and destination are located in the influence area of the same transit line do not have to make transfers. In these networks, there are either zero-transfer or one-transfer trips; the probability of the latter is equivalent to e_T , and the probability of direct trips is its complement.

Table A.2. Formulation of partial costs, commercial speed and occupancy constraint for grid network.

Partial cost	Grid network
L	$L = 2D^2/s$
V	$V = 4D^2[(1 - \phi)H_c + \phi H_p]/sH_cH_p$
M	$M = V/v_c$
v_c	$v_c = 1/[1/v_{FFS} + \tau/s + \tau'(1 + e_T)\Lambda/V]$
A	$A = s/w$
W	$W = H_c[\phi^2 D^2(1 + 3\phi + \rho(1 - \phi)) - 2\phi Ds(\rho + \phi) + \rho s^2(1 + \phi)]/2\phi^2 D^2(1 + \phi) + H_p(1 - \phi)(1 - \rho)/2(1 + \phi)$
E	$E = D[3 + 3\phi + 4\phi^2 - \rho(3 - \phi)]/6(1 + \phi)$
T	$T = E/v_c$
e_T	$e_T = 1 - s[2\phi D(\phi + \rho) - s\rho(1 + \phi)]/\phi^2(1 + \phi)D^2$
O	$O = (SF)(\Lambda s/4D)\max\{H_c[1 + 3\phi + \rho(1 - \phi)]/2\phi(1 + \phi); H_c(1 - \rho)/\phi; H_p(1 - \rho)/(1 + \phi)\}$

To estimate the occupancy, Figure 3 shows the most loaded points for each network structure, and at the same time, the paths followed by the users to account for how many cross each critical point. In a radial network, the maximum load is at the central stop. All the demand crosses this point, with the exception of those trips whose origin and destination belong to the influence area of the same line and the same quadrant. In a grid network, the most loaded point depends on the value of ϕ and the values of the central and peripheral headways. The distribution of passengers is assumed to be uniform among the different corridors taking into account the central corridors' prevalence due to their better headway of service.

On the other hand, the hybrid network's formulation is summarized in Table A.3. Its derivation follows the same procedure as in Daganzo (2010); however, the same aspects as previous networks are taken into account again. In addition, for this network structure, the distinction between central and peripheral areas is made from two different points of view: supply and demand. The former is due to the transit network structure, a grid at the central area and branched lines at the periphery, and the decision variable α determines the boundary between both. The latter is a consequence of the demand decentralization degree, where ϕ fixes the size of the attractant central area. For this reason, some user costs and capacity constraints are conditioned by the relationship between α and ϕ . If $\alpha > \phi$, all destinations are located at the gridded central area, and the users can reach their destinations with zero or one transfer. On the other hand, when $\alpha \leq \phi$, some destinations are located between the supply boundary αD and the demand boundary ϕD ; that is, at the branched lines. Therefore, some trips are completed by means of two transfers.

Table A.3. Formulation of partial costs, commercial speed and occupancy constraint for hybrid network.

Partial cost	Hybrid network
L	$L = D^2(1 + \alpha^2)/s$
V	$V = 2\alpha D^2(3 - \alpha)/sH$
M	$M = V/v_c$
v_c	$v_c = 1/[1/v_{FFS} + \tau/s + \tau'(1 + e_T)\Lambda/V]$
A	$A = s/w$
<i>If $\alpha > \phi$</i>	
W	$W = H[\phi^2 D^2[(2 + \alpha^3)(1 - \rho) + 3\alpha(1 + \rho) - 6\alpha\phi^2] - 3\phi Ds[\phi(1 + \alpha^2)(1 - \rho) + 2\alpha(\rho - \phi^2)] + 3\alpha s^2(1 - \phi^2)\rho]/6\alpha\phi^2(1 - \phi^2)D^2$

E	$E = D[(6\alpha + \phi^2 + \alpha^2\phi^2)(1 - \rho) + 8\alpha\phi(\rho - \phi^2)]/12\alpha(1 - \phi^2)$
T	$T = E/v_c$
e_T	$e_T = 1 - s[\phi^2 D(1 + \alpha^2)(1 - \rho) + 2\alpha\phi D(\rho - \phi^2) - \alpha s(1 - \phi^2)\rho]/\alpha\phi^2(1 - \phi^2)D^2$
O	$[H\Lambda s/8\alpha\phi(1 - \phi^2)D]\max\{2\alpha(1 - \rho)(2 - \phi^2 - \alpha^2); (1 - \rho)(\alpha + \phi)(2 - \alpha^2) + 2\alpha(\rho - \phi^2)\}$
<hr/>	
$If \alpha \leq \phi$	
<hr/>	
$W = \rho W_c + (1 - \rho)W_p$	
<hr/>	
W	$where:$ $W_c = H[D^2(\phi^4(8\phi + 3\alpha) + \alpha^3(3\alpha^2 - 2\phi^2) - \alpha^3(\phi + \alpha)) - 2Ds(\phi^3(2\phi + \alpha) + \alpha^3(2\alpha + \phi)) + 6\alpha^3s^2]/12\alpha\phi^4D^2$ $W_p = H[D(\phi^3(8\phi + 11\alpha) + \phi^2\alpha(4 + 3\alpha) - \alpha^3(\phi + \alpha)) - 2s(\phi^2(2\phi + 3\alpha) + \alpha^2(3\phi - 2\alpha))]/12\alpha\phi^2(\alpha + \phi)D$
E	$E = D[(\alpha^3(2\phi^2(1 + \rho) - \alpha^2\rho) + \phi^4(12\phi - 7\alpha))(1 + \phi) + 6\phi^4(1 - \rho)]/12\phi^4(1 + \phi)$
T	$T = E/v_c$
e_T	$e_T = 1 + [D(\phi^2 - \alpha^2) + 2\alpha s](\phi^2 - \rho\alpha^2)/2\phi^4D - [2\phi^3D - \rho\alpha^2s]s/\phi^4D^2$
O	$(H\Lambda/2\alpha\phi^2D)\max\{[D\alpha(\phi^2 - \alpha^2)(\phi^2 - \rho\alpha^2) + 4s(\phi^2\alpha^2(1 + \rho) + 3(\phi^4 - \rho\alpha^4))]/16\phi^2; s(\phi^2 - \alpha^2\rho)\}$

A.2. Direct trip-based network structure formulation

A more exhaustive explanation is required to prove this formulation. This scheme has its own characteristics. Following the explanations in Section 2.2, the attractant area is divided into concentric swaths of width d . The central one is served by a radial network, where all lines are gathered in the central point of the city and branched as the stop spacing s marks. The remainder swaths are served by different group of lines that collaborate to connect their respective swaths to the external area of the city, as yet unconnected. Figure 2 shows this situation and defines three important concepts for the succeeding proofs: swath-quadrant, external-quadrant and swath-corridor. Moreover, this figure shows that each line has two sections: (i) one that runs longitudinally through the swath-corridor, and (ii) one that runs through the branched itinerary of the central radial network. Each swath-corridor is identified by its distance from the city center $i \cdot d$, where $i=1,2,\dots,n$ and n , whose value is $(\phi D - d)/2d$, is the total number of swaths without taking into account the central one.

On the other hand, as Section 2.3 explains, two types of services are accepted: in headways or by schedules, depending on whether the headway of service is lower or higher than a fixed value H_s respectively. Therefore, two zones are identified: one around the city center and another external. The parameter ε_H , whose value is H_s/HD , fixes the boundary between both zones at a distance $\varepsilon_H D/2$ from the city center, a square cordon of side $\varepsilon_H D$.

Finally, imaginary square cordons are used to identify the trip's origin and destination. The parameter β , whose value varies between 0–1, defines these cordons by means of βD , which indicates the length of their sides, and $\beta D/2$, the distance from the city center. Vehicle flow is conserved in each cordon.

Result 1. The total length of the two-way infrastructure system is: $L = D^2/s + (\phi^2 D^2 - d^2)/d$.

Proof. The central swath is served by a radial network that, as known from Table A.1, has a length of D^2/s . For the remainder of the swaths, only the section of the lines that crosses longitudinally the swaths adds length, since the branched section is overlapped with the radial network. In the i -swath, the corridor is located at a distance $i \cdot d$ from the center, and forms a square cordon of side $2 \cdot i \cdot d$. Therefore, the total length of the corridor section is the summation for all the swaths $\sum_1^{(\phi D - d)/2d} 8id = (\phi^2 D^2 - d^2)/d$. \square

Result 2. The total vehicle-distance traveled per hour is: $V = [\phi D^2(6 + \phi) - d^2]/dH$.

Proof. The vehicle-kilometers traveled by the fleet that serves the central swath is $6D/H$ from Table A.1. For the remainder of the swaths, Figure 2 shows the distance traveled by one line. To explain this, we focus our attention on the lines of Figure 2a, which are gathered in the East swath-quadrant and are branched over the North and South external-quadrants. These lines run completely the East swath-quadrant, and on average half of both adjacent ones; that is, half of the corridor's length. For the i -swath, this distance is $4 \cdot i \cdot d$.

In the branched section, the lines cross the North external-quadrant vertically from its North swath-quadrant to the city boundary. This distance is $(D/2 - id)$. At the same time as these lines cross vertically that external-quadrant, some horizontal displacements exist due to the route branching. One line runs approximately $s/2$ units of distance horizontally per each s unit of distance vertically. The number of times that a line travels s units of distance vertically is $(D/2 - id)/s$; therefore, the total horizontal distance is $(s/2)(D/2 - id)/s$. This distance is doubled, since the same length is traveled in the South quadrant; that is, $2[(D/2 - id) + (s/2)(D/2 - id)/s] = 3(D - 2id)/2$.

As Section 2.2 explains, the number of vehicles that cover these distances per each swath-quadrant in the i -swath is $1/H$. The resultant kilometers traveled is multiplied by two due to the bidirectional service of the lines, and by four because there are four groups of lines per swath, giving $2 \cdot 4 \cdot [4id + 3(D - 2id)/2](1/H) = 4(3D + 2id)/H$. Adding the vehicle-distance traveled in each swath, the total vehicle kilometers in one hour V is $6D/H + \sum_1^{(\phi D - d)/2d} 4(3D + 2id)/H = [\phi D^2(6 + \phi) - d^2]/dH$. \square

Result 3. The expected commercial speed during rush hour is: $v_c = 1/[1/v + \tau/s + \tau' \Lambda/V]$.

Proof. As with Daganzo (2010), the total time needed for a vehicle to travel a distance is the sum of the time that one vehicle would need to complete that distance if it did not make stops, the time spent to make those stops and the time that the users need to get on and get off the vehicle at each stop. However, there are no transfers in this case, meaning fewer boardings and alightings. \square

Before deriving the user costs, different types of trips are distinguished by their origin's location. There are two groups: (a.1) those whose origins are located at the central attractant area, which generates and attracts demand; and (a.2) those whose origins are peripheral, located at the external area where there is only trip generation. The probability of (a.1) and (a.2) are ρ and $(1-\rho)$ respectively. If the demand is uniformly generated over the whole city, that is, $f_d=1$, these probabilities are ϕ^2 and $1 - \phi^2$.

Result 4. The expected walking time at the origin and destination (access and egress) is: $A = (3s + d)/4w$.

Proof. As Section 2.3 explains, the users take the transit vehicle at the closest stop. The influence area of one stop is a centered square of side s . The average access distance is $s/2$, zero for the closest origin, located at the same stop, and s for the furthest origin, located at one square's vertex. On the contrary, the users alight at stops of a swath-corridor to avoid transfers, although these stops are not always the closest. These stops have a rectangular influence area of side s in the longitudinal direction and side d in the perpendicular direction. Following the same reasoning as for the access, the average egress distance is $(s + d)/4$. In some trips, this behavior is reversed, users go to a swath-corridor stop at origin, and they alight at the closest stop to their destinations. The sum of access and egress distances divided by the pedestrian speed gives the expected walking time: $[s/2 + (s + d)/4]/w = (3s + d)/4w$. \square

Result 5. The expected waiting time per user at a stop is:

$$\begin{aligned} \text{if } \varepsilon_H \geq 1 & \quad W = H[5(1 + \phi + \phi^2) - \rho(5 - \phi - \phi^2)]D/15(1 + \phi)s \\ \text{if } \phi \leq \varepsilon_H < 1 & \quad W = \left[h_s(1 - \varepsilon_H^2) + H(2f_s(1 - \varepsilon_H^3) + (\varepsilon_H^3 - \phi^3)) \right] D/3s \left[(1 - \rho)/(1 - \phi^2) + 2\phi DH\rho/5s \right] \\ \text{if } \varepsilon_H < \phi & \quad W = [h_s + 2f_s(1 + \phi + \phi^2)DH/3(1 + \phi)s](1 - \rho) \\ & \quad + [h_s(\phi^4 - \varepsilon_H^4) + 2H(2f_s(\phi^5 - \varepsilon_H^5) + \varepsilon_H^5)D/5s]\rho/\phi^4 \end{aligned}$$

Proof. First, we define the headway of service for the different network's stops. In the central radial network, the headway of service increases from the city center, whose value is H , with a slope $\beta D/s$. In a cordon located at a distance $\beta D/2$, the number of points where the transit lines cross one side of this cordon is $\beta D/s$. Moreover, in each cordon, the flow of vehicles is constant, and its value is $1/H$ for one of its sides. This flow is uniformly distributed among the different points crossed. Therefore, the headway of service at stops that belong to the cordon βD is the inverse of the ratio between that number of vehicles and the number of points crossed: $1/[(1/H)/(\beta D/s)] = (\beta D/s)H$. The same happens with the groups of lines that serve the remainder of the swaths. The lines of the same group are gathered with a joint headway H in the swath-corridor, and follow the same branching as the radial network. As a consequence,

one stop located at distance $\beta D/2$ from the city center is connected to each inner swath by a headway of service $(\beta D/s)H$.

On the other hand, the waiting time at stops depends on the type of service with which the transit network works. If it works in headways, it is assumed that the waiting time is half the headway. On the contrary, the waiting time is composed of two terms, e.g. Tirachini et al. (2010). These are a constant time h_s independent of the headway, and a variable term as an opportunity cost. This variable term is the result of multiplying a fixed factor f_s by the headway. There is a cut-off headway H_{h-s} that indicates how each stop works: in headways if the headway of service at stop is lower than H_{h-s} ; otherwise, by schedules.

In addition, the parameter ε_H , previously presented, defines how the transit network works. At cordon $\varepsilon_H D$, the headway is $(\varepsilon_H D/s)H$, which is equalized to H_{h-s} . Therefore, the value of ε_H is H_{h-s}/HD . Given its value, we can distinguish between three scenarios: (b.1) $\varepsilon_H \geq 1$, the whole system works in headways, (b.2) $\phi \leq \varepsilon_H < 1$, the outer section of the periphery works by schedules, and (b.3) $\varepsilon_H < \phi$, only an inner zone of the central attractant area works in headways. Moreover, ε_H and ϕ define different areas and their probabilities in each scenario. In the case of (b.1), there are only two areas: the same as (a.1) and (a.2); for the scenario (b.2) the peripheral area (a.2) is divided into one outer and one inner zone, whose probabilities are $(1 - \varepsilon_H^2)(1 - \rho)/(1 - \phi^2)$ and $(\varepsilon_H^2 - \phi^2)(1 - \rho)/(1 - \phi^2)$ respectively, and the central area is equivalent to (a.1); and finally, in the case (b.3), its periphery matches with (a.2) and the central area is partitioned in two zones, whose probabilities are $(\phi^2 - \varepsilon_H^2)\rho/\phi^2$ for the outer and $\varepsilon_H^2\rho/\phi^2$ for the inner.

The expected waiting time for scenario (b.1) is now derived. At the periphery, the average headway is the weighted headway of each cordon βD , taking into account that the probability density function of β is triangular; that is, $2\beta/(1 - \phi^2)$. The result is $H_p = \int_{\phi}^1 [(2\beta/(1 - \phi^2))(\beta D/s)H] d\beta = 2(1 + \phi + \phi^2)DH/3(1 + \phi)s$. At the central area, the headway for each trip is the highest between the headways of its origin (β_o) and destination (β_d) stops, where the p.d.f. of β_o and β_d is $2\beta_i/\phi^2$. Therefore, $H_c = \int_0^{\phi} (2\beta_1/\phi^2) [\int_0^{\beta_1} [(2\beta_2/\phi^2)(\beta_1 D/s)H] d\beta_2 + \int_{\beta_1}^{\phi} [(2\beta_2/\phi^2)(\beta_2 D/s)H] d\beta_2] d\beta_1 = 4\phi DH/5s$. All this leads to obtaining the expected waiting time: $W = [H_c\rho + H_p(1 - \rho)]/2 = H[5(1 + \phi + \phi^2) - \rho(5 - \phi - \phi^2)]D/15(1 + \phi)s$.

For the other scenarios, when ε_H is lower than 1, the same process is followed to derive the expected waiting time, only with the exception that some parts of the city work by schedules. If $\varepsilon_H \geq \phi$, this only happens in the outer periphery, and outer and inner average peripheral headways are distinguished related to the trip's origin. These headways are $H_p^o = \int_{\varepsilon_H}^1 [(2\beta/(1 - \varepsilon_H^2))(\beta D/s)H] d\beta = 2(1 + \varepsilon_H + \varepsilon_H^2)DH/3(1 + \varepsilon_H)s$ and $H_p^i = \int_{\phi}^{\varepsilon_H} [(2\beta/(\varepsilon_H^2 - \phi^2))(\beta D/s)H] d\beta = 2(\varepsilon_H^2 + \varepsilon_H\phi + \phi^2)DH/3(\varepsilon_H + \phi)s$ respectively. On the contrary, all the periphery works by schedules, such as the most external region of the attractant central area. Therefore, in this case, the central average headway is divided into three types of trip: (i) origin and destination in the outer region $H_c^o = \int_{\varepsilon_H}^{\phi} (2\beta_1/(\phi^2 - \varepsilon_H^2)) [\int_{\varepsilon_H}^{\beta_1} [(2\beta_2/(\phi^2 - \varepsilon_H^2))(\beta_1 D/s)H] d\beta_2 + \int_{\beta_1}^{\phi} [(2\beta_2/(\phi^2 - \varepsilon_H^2))(\beta_2 D/s)H] d\beta_2] d\beta_1 = 4(2\varepsilon_H^3 + 4\varepsilon_H^2\phi + 6\varepsilon_H\phi^2 + 3\phi^3)DH/15(\varepsilon_H + \phi)^2s$, (ii) origin or destination in the outer region and the other in the inner $H_c^{i,o} = \int_{\varepsilon_H}^{\phi} [(2\beta/(\phi^2 - \varepsilon_H^2))(\beta D/s)H] d\beta = 2(\phi^2 + \phi\varepsilon_H + \varepsilon_H^2)DH/3(\varepsilon_H + \phi)s$, and (iii) both in the inner area $H_c^i = \int_0^{\varepsilon_H} (2\beta_1/\varepsilon_H^2) [\int_0^{\beta_1} [(2\beta_2/\varepsilon_H^2)(\beta_1 D/s)H] d\beta_2 + \int_{\beta_1}^{\varepsilon_H} [(2\beta_2/\varepsilon_H^2)(\beta_2 D/s)H] d\beta_2] d\beta_1 = 4\varepsilon_H DH/5s$. The resulting expected waiting times for each scenario are:

$$\begin{aligned} \text{if } \phi \leq \varepsilon_H < 1 \quad W &= [(h_s + f_s H_p^o)(1 - \varepsilon_H^2) + H_p^i(\varepsilon_H^2 - \phi^2)/2](1 - \rho)/(1 - \phi^2) + \rho H_c \phi^2/2 = [h_s(1 - \varepsilon_H^2) + \\ &\quad H(2f_s(1 - \varepsilon_H^3) + (\varepsilon_H^3 - \phi^3))D/3s](1 - \rho)/(1 - \phi^2) + 2\phi DH\rho/5s \\ \text{if } \varepsilon_H < \phi \quad W &= (h_s + f_s H_p)(1 - \rho) + [(h_s + f_s H_c^o)(\phi^2 - \varepsilon_H^2)^2 + (h_s + f_s H_c^{i,o})2\varepsilon_H^2(\phi^2 - \varepsilon_H^2) + \\ &\quad H_c^i \varepsilon_H^4/2]\rho/\phi^4 = [h_s + 2f_s(1 + \phi + \phi^2)DH/3(1 + \phi)s](1 - \rho) + [h_s(\phi^4 - \varepsilon_H^4) + \\ &\quad 2H(2f_s(\phi^5 - \varepsilon_H^5) + \varepsilon_H^5)D/5s]\rho/\phi^4 \quad \square \end{aligned}$$

Result 6. The expected in-vehicle travel distance per trip is: $E = [15(1 - \rho) + 2\phi(1 + \phi)(10 + \rho)]D/30(1 + \phi)$.

Proof. The length of all trips is decomposed into two sections. One runs over a branched route, and the other partially a swath-corridor. The former has the same behavior as in Daganzo (2010): the traveled

distance between an origin at the cordon of side $\beta_o D$ and a destination at the cordon of side $\beta_d D$ is one perpendicular $|\beta_o - \beta_d|D/2$ and one transverse $|\beta_o - \beta_d|D/4$ due to the route branching. Regarding the second section, a swath-corridor located at a distance $\beta D/2$ from the city center shapes a square cordon of side βD . Therefore, the expected distance in the swath-corridor is the average distance between two random points of that cordon by the shortest possible path; that is, βD units of distance.

First, we compute the expected trip length when the origin is located at the periphery. In this case, there is a section of the trip in that periphery and another in the central attractant area. As the parameter β has a triangular density function $2\beta/(1-\phi^2)$, this length is $E_p = \int_{\phi}^1 [(2\beta_o/(1-\phi^2))(3(\beta_o - \phi)D/4)] d\beta_o + \int_0^{\phi} [(2\beta_d/\phi^2)(3(\phi - \beta_d)D/4)] d\beta_d = (3 + 4\phi + 4\phi^2)D/6(1 + \phi)$. On the other hand, when the origin and destination are in the central attractant area, the average distance is $E_c = \int_0^{\phi} \int_0^{\phi} (2\beta_d/\phi^2)(2\beta_o/\phi^2)[3|\beta_o - \beta_d|D/4 + \min(\beta_o, \beta_d)D] d\beta_o d\beta_d = 11\phi D/15$. Weighting peripheral and central distances by $(1-\rho)$ and ρ respectively, the expected traveled distance is $E = [15(1-\rho) + 2\phi(1+\phi)(10+\rho)]D/30(1+\phi)$. \square

Result 7. *The expected vehicle occupancy on the critical load point during rush hour is:*

$$\text{if } (2d\rho + \sqrt{3\phi^2 D^2 \rho + d^2 \rho^2})/6\rho \leq \phi D/2 \quad O = (SF)O_i$$

$$\text{if } (2d\rho + \sqrt{3\phi^2 D^2 \rho + d^2 \rho^2})/6\rho > \phi D/2 \quad O = (SF)O_e$$

Proof. The highest vehicle occupancy in the transit network takes place on the swath-corridors. As explained in Section 2.2, all the demand from an external-quadrant whose destination belongs to the same adjacent swath-quadrant is carried by a group of lines that are gathered on the respective swath-corridor. For a corridor located at a distance $i \cdot d$, the generated demand from its external-quadrant is $(1-\rho)\Lambda/4 + \rho\Lambda[\phi D/2 - (id - d/2)][\phi D + (2id - d)]/2\phi^2 D^2$, and the probability that this demand is attracted by that swath-quadrant is $2id^2/\phi^2 D^2$. As Figure 3 shows, these lines also connect that external-quadrant to the same and the opposed swath-quadrants. For some users, these lines serve one, the other or a percentage of both. On average, these lines connect half for each; that is, an additional swath-quadrant, so that the demand carried is multiplied by two. This demand is allocated among those transit vehicles that serve the corresponding lines, whose flow is $1/H$. Then, the vehicle occupancy on the corridor i is $H\Lambda id^2 [\rho(\phi^2 D^2 - d^2(1 + 4i(i-1))) + (1-\rho)\phi^2 D^2]/\phi^4 D^4$.

As the previous expression shows, not all of the corridors carry the same passenger load: it depends on the size of each swath and the size of the external-quadrant that each of them serves. The first determinant increases with the distance from the city center, but the second decreases. The consequence is that the corridors' passenger load shows a concave behavior with distance from the center, i.e., the maximum load is not located at extremes, but at an intermediate distance from the center. The value of i that maximizes the vehicle occupancy is $(2d\rho + \sqrt{3\phi^2 D^2 \rho + d^2 \rho^2})/6\rho$; multiplied by d it gives the distance from the city center where the corridor with maximum load is located. If this distance is shorter than $\phi D/2$, there is a corridor around this position. Then, the maximum occupancy is $O_i = H\Lambda d[d\rho(9\phi^2 D^2 - \rho d^2) + \rho^{1/2}(3D^2 + d^2)^{3/2}]/27\phi^4 D^4 \rho$. However, if that position is situated in the periphery where there are no corridors, the corridor with the highest load is around the boundary of the central attractant area. This happens because the occupancy function has an increasing monotonous behavior with respect to the corridor's position defined by i . In this case, the maximum occupancy is $O_e = H\Lambda d(\phi D - d)[\phi^2 D^2(1-\rho) + 4d\rho(\phi D - d)]/2\phi^4 D^4$.

In addition, the rush hour actually has its own peaks; for this reason, the occupancy is exacerbated by a safety factor (SF). \square

Finally, the parameter e_T is null in this network structure since all the trips are direct, and the number of vehicles that serve the network and the expected in-vehicle time for users are derived as in Daganzo (2010) as $M = V/v_c$ and $T = E/v_c$, respectively.

References

- Aldaihani, M.M., Quadrioglio, L., Dessouky, M.M., Hall, R., 2004. Network design for a grid hybrid transit service. *Transportation Research Part A* 38 (7), 511–530.
- Anas, A., Arnott, R., Small, K., 1998. Urban spatial structure. *Journal of Economic Literature* 36 (3), 1426–1464.
- Badia, H., Estrada, M., Robusté, F., 2014. Competitive transit network design in cities with radial street patterns. *Transportation Research Part B* 59, 161–181.
- Bertaud, A., 2004. The Spatial Organization of Cities: Deliberate Outcome or Unforeseen Consequence?. *IURD Working Paper Series* WP-2004-01, Institute of Urban and Regional Development, Berkeley, CA.
- Brown, J.R., Thompson, G. L., 2008. Examining the Influence of Multidestination Service Orientation on Transit Service Productivity: A Multivariate Analysis. *Transportation* 35 (2), 237–252.
- Brown, J.R., Thompson, G.L., 2012. Should Transit Serve the CBD or a Diverse Array of Destinations? A Case Study Comparison of Two Transit Systems. *Journal of Public Transportation* 15 (1), 1–18.
- Chen, H., Gu, W., Cassidy, M.J., Daganzo, C.F., 2015. Optimal transit service atop ring-radial and grid street networks: A continuum approximation design method and comparisons. *Transportation Research Part B: Methodological* 81, 755–774.
- Craig, S.G., Ng, P.T., 2001. Using quantile smoothing splines to identify employment subcenters in a multicentric urban area. *Journal of Urban Economics* 49 (1), 100–120.
- Daganzo, C.F., 2010. Structure of competitive transit networks. *Transportation Research Part B* 44 (4), 434–446.
- Estrada, M., Roca-Riu, M., Badia, H., Robusté, F., Daganzo, C.F., 2011. Design and implementation of efficient transit networks: Procedure, case study and validity test. *Transportation Research Part A* 45 (9), 935–950.
- Heddebaut, O., Finn, B., Rabuel, S., Rambaud, F., 2010. The European Bus with High Level of Service (BHLS): Concept and Practice. *Built Environment*, Lloyd Wright editor “*Bus Rapid Transit: A public renaissance*” 36 (3), 307–316.
- Jansson, J.O., 1980. A simple bus line model for optimization of service frequency and bus size. *Journal of Transport Economics and Policy* 14 (1), 53–80.
- Jara-Díaz, S. R., Gschwender, A., 2003. From the Single Line Model to the Spatial Structure of Transit Services: Corridors or Direct?. *Journal of Transport Economics and Policy* 37 (2), 261–277.
- Lee, B., 2007. “Edge” or “edgeless” cities? Urban spatial structure in US metropolitan areas, 1980 to 2000. *Journal of Regional Science* 47 (3), 479–515.
- Li, Z., Chen, Y., Wang, Y., 2013. Optimal Density of Radial Major Roads in a Two-dimensional Monocentric City with Endogenous Residential Distribution and Housing Prices. *Regional Science and Urban Economics* 43 (6), 927–937.
- Mees, P., 2000. A very public solution: transport in the dispersed city. Melbourne University Press, Melbourne.
- Newell, G.F., 1979. Some issues relating to the optimal design of bus routes. *Transportation Science* 13 (1), 20–35.
- Nielsen, G., Nelson, J.D., Mulley, C., Tegner, G., Lind, G., Lange, T., 2005. Public transport - Planning the networks. HiTrans Best Practice Guide Vol. 2, Stavanger, Norway.
- Nourbakhsh, S.M., Ouyang, Y., 2012. A structured flexible transit system for low demand areas. *Transportation Research Part B* 46 (1), 204–206.
- Pereira, R.H.M., Nadalin, V., Monasterio, L., Albuquerque, P.H.M., 2013. Urban centrality: a simple index. *Geographical Analysis*, 45 (1), 77–89.
- Rodrigue, J.P., Comtois, C., Slack, B., 2006. The geography of transport systems. Routledge, London.
- Schimke, P., 1997. Understanding Differences in Public Transit: Comparison of Boston and Toronto. *Transportation Research Record* 1604 (1), 9–17.
- Thompson, G.L., 1977. Planning considerations for alternative transit route structures. *Journal of American Institute of Planners* 43 (2), 158–168.
- Thompson, G.L., Matoff, T.G., 2003. Keeping up with Joneses: planning for transit in decentralizing regions. *Journal of the American Planning Association* 69 (3), 296–312.
- Tirachini, A., Hensher, D.A., Jara-Díaz, S.R., 2010. Comparing operator and users costs of light rail, heavy rail and bus rapid transit over a radial public transport network. *Research in Transportation Economics* 29 (1), 231–242.
- Tsekeris, T., Geroliminis, N., 2013. City Size, Network Structure and Traffic congestion. *Journal of Urban Economics* 76, 1–14.
- Vaughan, R., 1986. Optimum polar networks for an urban bus system with a many-to-many travel demand. *Transportation Research Part B* 20 (3), 215–224.
- Vaughan, R., 1987. Urban spatial traffic pattern. Pion Ltd., London.